

APPLICATION OF

THOMAS CHRISTOPHER DYER

APPLICATION FOR LETTERS PATENT OF THE UNITED STATES

**METHOD AND SYSTEM FOR DISPLAYING RELATED COMPONENTS OF A MEDIA  
STREAM THAT HAS BEEN TRANSMITTED OVER A COMPUTER NETWORK**

James J. DeCarlo  
Registration No. 36,120  
Attorney for Applicant  
STROOCK & STROOCK & LAVAN LLP  
180 Maiden Lane  
New York, New York 10038  
(212) 806-5400

Atty. Docket No.: 694231/0017

"Express Mail" mailing label No.

Number CL 1235628530 US

Date of Deposit 5/16/01

I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to the Commissioner for Patents,  
Washington, D.C. 20231.

Name

Signature

Jennifer Bartolo  
Jennifer Bartolo

## **FIELD OF THE INVENTION**

This invention relates to a system and apparatus for displaying related components of a media stream that has been transmitted over a computer network. More specifically, the invention relates to delivering closed caption and other information that may be included in the television broadcast to an audio-visual media player.

## **BACKGROUND OF THE INVENTION**

In order to display video on a television monitor, an electron gun inside the monitor's cathode ray tube moves across the screen line by line as it scans its way down to the bottom of the screen. Video transmission is delayed while the electron gun moves back up to the top of the tube to display the next image. During this delay, called the vertical blanking interval ("VBI"), other data may be delivered to the display. For example, closed captioning text, stock market tickers and news and weather updates are often embedded in television broadcast signals. In particular, closed caption data is typically embedded in the broadcast at line 21 of the VBI. This system allows the more data to be delivered to the television and, therefore, provides for more efficient operation.

Thanks to advances in streaming media technology, television signals can also be transmitted and viewed on the Worldwide Web. Streaming media generally includes a digitized set of images and sounds that are sent over the Internet and played for the viewer at an end user station in sequence. Significantly, the image and sounds are played as they arrive. The use of streaming media, therefore, allows an Internet user to avoid having to download an entire file before being able to play any audio or video. This often saves a substantial amount of time and eliminates considerable frustration.

While closed caption and other information is easily displayed on a television set, currently available streaming media players cannot always display such information in an Internet browser. For example, closed caption data is typically not delivered in a format that can be used by a streaming media player. More specifically, closed caption data is typically delivered as single letters or in segments of words rather than as entire words or phrases. Thus, it must be displayed by a device that can simultaneously display information that is received at different

times. In contrast, audio-visual media players are set up to receive information as a stream, and to continuously replace existing data with new data as it is received. Currently available methods, therefore, cannot typically be used to display closed caption text using an audio-visual media player.

### **SUMMARY OF THE INVENTION**

The invention is generally directed to a system and apparatus for displaying components of a composite information stream that has been transmitted over a computer network. More specifically, the invention relates to displaying closed caption and other information that can be embedded in television signals in a Worldwide Web browser.

In one embodiment, the invention includes a system for simultaneously displaying multiple components of a composite information stream, which includes a storage device that has been configured to communicate with a decoder (or other device that is capable of parsing the composite stream into components) and with an information display system. The storage device is capable of receiving at least one component of the composite information stream from the information extractor, and of transmitting the composite information stream component to the information display system in response to an information release signal that is embedded in the information stream component.

In another embodiment, the invention includes an information storage device, such as a buffer or other computer storage device, that has been configured to communicate with a decoder and with an information display system. The buffer is capable of receiving at least a portion of the text information from the decoder, and is capable of transmitting the received text information to the information display system in response to an information release signal that is contained in the text information.

In accordance with an embodiment of the invention, a method of simultaneously displaying multiple types of information includes parsing multiple types of related information into a plurality of single information type components and collecting at least one of the single information type components in the storage device. At least one of the single information type components includes an information release signal. In accordance with the method, the content of the information storage device can be transferred to an encoder and forwarded over the

Internet to an information display system in response to the information release signal. The information storage device content and at least one other corresponding single information type component are simultaneously presented on the display.

At least one embodiment of the invention includes a system for delivering closed caption information with a television broadcast. The closed caption data is delivered to the buffer, and retained there until a release signal is generated. When the release signal is detected, the entire content of the buffer is forwarded to an encoder and transmitted over the Internet to an audio-visual media player for display on a video monitor.

Other embodiments of the present invention and features thereof will become apparent from the following detailed description, considered in conjunction with the accompanying drawing figures.

### **BRIEF DESCRIPTION OF THE DRAWINGS**

FIG. 1 is a schematic illustration that depicts an example of a commonly known system that is capable of simultaneously displaying related components of a television signal on a video monitor or similar device;

FIG. 2 is a system diagram illustrating one way that the present invention can be implemented in a system for displaying multiple components of a composite television signal on the Internet; and.

FIG. 3 is a flow diagram with exemplary steps that are currently followed to deliver audio and video data with corresponding closed caption data to viewer on a television monitor.

FIG. 4 is a flow diagram that illustrates steps that can be followed to display multiple components of a television signal, such as closed-caption data with corresponding audio and video to a viewer over the Internet.

FIG. 5 is a schematic illustration of a television set with video and closed caption data displayed in the viewing area..

FIG. 6 is a schematic illustration of an end user processor with video and closed caption data displayed in the viewing area in accordance with an embodiment of the invention.

While the present invention will be described in connection with certain embodiments thereof, it is to be understood that the invention is not limited to those embodiments. On the contrary, it is intended to cover all alternatives, modifications, and equivalents as may be included within the spirit and scope of the invention as defined by the appended claims.

### **DETAILED DESCRIPTION OF EMBODIMENTS OF THE INVENTION**

Referring now to the drawings which are provided to describe embodiments of the invention and not for limitation, FIG. 1 contains a schematic illustration of a system that can simultaneously display related components of a television signal on a television set, video monitor or similar device. Television signals can generally be described as composite information. That is, they contain multiple components of distinct types of information (e.g. audio, video, closed caption, etc.). Generally speaking, a system for displaying such information includes a receiver 12 to which the television signals are delivered from a television broadcast station, cable television provider, digital satellite system operator or other comparable source. Such television signals generally include audio and video information, and they may also include other types of data that relates to the program being broadcast. As shown, receiver 12 is linked to a decoder 14, which 14 may be located inside a set top box, or in a television set, video monitor or other device upon which television images will be displayed. While such a configuration is very common, it is also possible to provide a system 10 with a decoder 14 that is located at the broadcast station or in a remote location.

One readily available decoder 14 is a VBI decoder, which can be used to extract data that is delivered to the system in the vertical blanking interval of the television broadcast as described above. The VBI decoder can extract the desired portions of the television signal, and allow the remaining components to pass through as appropriate. Thus, separate components can be delivered to the speakers (for the audio component, not shown) and to designated regions 16 and 18 of the display for listening and viewing.

While a system such as that described with reference to FIG. 1, is adequate for displaying information on a television set, difficulties arise when it is desired to display television signals in other ways. For example, the growth in popularity of the Internet makes it

desirable to deliver television signals to over the Worldwide Web. While the development of streaming media technology has made it possible to deliver the audio and video portions of the television signal to Internet users, audio-visual media players that are currently available cannot properly process all of the other components of the signals.

For example, media players are currently not configured to display closed caption data over the Internet. In a live video broadcast, this type of data is generally delivered to a television screen a single letter at a time or in small groups of letters. This is because the text must first be transcribed by a human operator, which obviously limits the rate at which the transcription can take place. Thus, entire sentences must be delivered to the display a single letter at a time. This is acceptable for delivery of data to a television set, because each letter can continue to be displayed on the screen until the entire word or sentence in which it is to be included is delivered to the screen and displayed. Also, slow delivery of the text to the television screen allows the viewer to be able to keep up with the text as it scrolls across the display region.

However, currently available Internet and other audio-visual media players are configured differently. Media players continuously replace existing data with new data as it is received. Thus, the first word of a sentence would be replaced by the second word, which would be replaced by the third word, etc., allowing only single words of the sentence to be displayed at a time. In order to transmit an entire sentence or more to an audio-visual media player, the system would have to send the first word and then replace it with both the first and second words, which would then be replaced by the first three words, and so on. Thus, it would be impossible to display entire sentences without including a large amount of unwanted data in the data stream. Known methods are, therefore, typically not used to deliver closed caption text to an audio-visual media player with corresponding audio and video.

Referring to FIG. 2, according to the present invention, a storage device 20 is configured to receive the data that has been extracted from decoder 14 and to transmit it over a computer network such as the Internet, for display on a monitor or other appropriate output device that is linked to end user processor 26. An end user processor 26 may be a personal computer, personal digital assistant, Internet appliance or other now known or yet to be known device for receiving and delivering humanly perceptible digital information over a computer network. End user processors 26 typically include an Internet browser, such as Internet Explorer

or Netscape Navigator, and a streaming media player such as Microsoft Windows Media Player, Real Networks' Real Player or a similar device. In one embodiment, storage device 20 is a buffer, which is commonly used in a computer system for the temporary storage of data. It should be noted that other storage devices, such as those that provide for the permanent storage of data could also be used. While the illustration shows one storage device 20 linked to a single decoder 14, it is to be understood that multiple decoders 14 and/or storage devices 20 could be provided. For example, multiple storage devices 20 could be linked to a single decoder 14 to store a large amount of a single component of extracted data. Or, a single decoder 14 could extract multiple types of data, and a separate storage device 20 could be provided for each type of extracted data. Another system could include multiple decoders 14, each of which extracts a different type of data from the television signals, and a separate storage device could be linked to each decoder 14.

As described earlier, television signals are delivered to receiver 12, and one or more components can be extracted by decoder 14. In one an embodiment of the invention, a component of the extracted information contains closed caption text data. This closed caption text is delivered to storage device 20, and is retained. When an appropriate signal is detected, the entire contents of storage device 20 are then transmitted, along with the components that are passed through decoder 14 without processing, to a media encoder 28. Encoder 28 merges the data to cause the extracted components to be delivered to the viewer simultaneously with the corresponding unprocessed components, and the merged data stream is delivered to streaming media server 22. End users may access streaming media and various other forms of content by viewing web sites at end user processors 26, and selecting various links.

FIG. 3 is a block diagram that illustrates the steps for delivering multiple television signal components, such as audio, video and closed-caption data, to a television set or video monitor. As shown, television signals are delivered from receiver 12 to VBI decoder 14 at block 102, which extracts the relevant data at block 104. Unextracted portions of the signal are passed to television receiver 12 as indicated in block 18, while the extracted data is passed to an appropriate encoder as shown in block 106. The extracted data is processed, and delivered to the display, along with the unextracted portions of the signal as indicated in block 110.

One way in which embodiments of the invention may be incorporated into a system such as that described above as illustrated in FIG. 4. In this example, television signals that include closed caption text and corresponding audio and video signals, may be delivered over a computer network such as the Internet and displayed on a video monitor. As before, television signals are transmitted from receiver 12 to VBI decoder 14 as shown in block 202, which extracts the closed caption data as shown in block 204. As stated earlier, television signals could include other types of data, and while the invention is described with reference to closed caption data, it is not limited to such use.

After the closed caption data has been extracted, the remaining components of the television signal pass to media player encode r28 as indicated in block 212. The extracted, closed caption data is delivered to a storage device as shown in block 206. As noted earlier, in one embodiment, the extracted data is stored in a buffer. In other embodiments, the extracted data may be stored in a permanent memory. In one embodiment of the invention, the extracted component (or components) is retained in storage device 20 until a release signal is detected as shown at block 208. This signal will typically be generated when an entire word, sentence or other designated set of information has been collected in storage device 20. In one embodiment of the invention, an ASCII "line feed" character (i.e. ASCII character 10) serves as the signal that causes the information to be released from storage. Such a character is often included in closed caption text at the end of a line, sentence or paragraph.

With continued reference to FIG. 4, when the release signal is detected at block 208, all of the information that has been collected in storage device 20 is transmitted to media encoder 28 as shown in block 210. Corresponding components of the television signal - the extracted closed caption data and the audio and video data to which it relates - are then transmitted together from encoder 28 to streaming media server 22 as indicated in block 214. The components are delivered to media players at end user processors as shown in block 216 as they are requested by users.

It should be noted that the placement of blocks 210 and 212 adjacent to each other in FIG. 3 does not imply that the invention requires data to be delivered to the media player simultaneously. The steps will typically be performed in the manner that will cause the audio, video, closed caption and other components to be properly encoded for delivery to the streaming



media server in a way that will cause the various components of the associated data to be played by the audio-visual player at the same time. Thus, any differences in processing speeds and transmission rates should be taken into account in order to designate any order or timing for performing steps 210 and 212.

In one embodiment of the invention, closed caption information is transmitted from storage device 20 as a text script command. As explained earlier, the text script command is transmitted to media encoder 28 at block 210 and sent to the streaming media server 22 as indicated in block 214. Closed caption information can be provided in several different types of script commands. In one embodiment, it may be delivered as ordinary text, which is displayed in a designated region 18 of the display. In another embodiment, the closed caption information may be provided as a Universal Resource Locator (URL) that is launched by the media player in a browser window (i.e. a "Slide Flip"). In still another embodiment, a customized script command may be provided. In such a case, the HTML file usually includes custom script code, which can be interpreted by the media player.

The functions that are illustrated in the various blocks of FIG. 4 can be incorporated into a visual basic application, or similar program. While these embodiments of the invention have been illustrated as being suitable for display on end user processors 26 that are connected to video monitors, it is to be understood that numerous display devices including, but not limited to, laptop computers and handheld personal digital assistants can be used. It should also be noted that the television signal components can be displayed on different devices if desired. Thus, the invention could be incorporated into a system that displays only the closed caption component, and therefore enables a user to read the information that has been included in the television signal, when access to the audio and video components cannot be provided. This information could also be displayed on a marquee or other large device, in order to provide useful information to those who are able to view it.

As indicated earlier, television systems which simultaneously display video data with corresponding closed caption data are well known. In these systems, video signals are typically displayed in a substantial region 16 of the viewing area while closed caption data is displayed in a smaller region 18, as indicated in FIG. 5. Because television screens are relatively large - often on the order of 27 inches in diameter to 35 inches or more - closed caption data that

is displayed in the relatively small region 18 is usually still easily interpreted by the viewer. In contrast, a video monitor with a diameter of only 21 inches is considered relatively large. The size of a video monitor is largely constrained by the more stringent resolution requirements for computer these devices. Further, streaming video screens only occupy a portion of the monitor on which they are displayed. For example, streaming media video screen resolutions on the order of 160x120 pixels, 176x132 pixels, 240x180 pixels and 320x420 pixels are quite common. It is not uncommon, on the other hand, to find video monitors that have resolutions on the order of 1024x768 pixels, 1280x1024 pixels or even 1600x1200 pixels. Thus, a streaming video window is quite small compared to that of a video monitor.

In any event, it would be much more difficult for a viewer to read closed caption data if it were displayed in a region that corresponds to the size of region 18 of FIG. 5. Further, video data that is delivered to end users via Internet web sites usually occupies less than the entire viewing area. This is typically because other portions of the site must still be displayed. Thus, it would be virtually impossible for a viewer to read closed caption data that were provided with corresponding video if it were delivered over the Internet in the same manner that it is delivered in a television system.

Turning to FIG. 6, the present invention provides an additional advantage, in that it delivers closed caption data to a closed caption region 18 that is specifically configured for a media player, rather than a television system. As shown, a media player often includes a video region 16 that is relatively small compared to that of a television screen (FIG. 5). However, the closed caption region 18 of the player is typically relatively large, and often lies outside of video region 16. Thus, it is much easier for a viewer to read closed caption data when it is delivered to a video monitor over the Internet in accordance with the invention.

In addition, television systems that display caption data typically include some mechanism, such as a button on the remote control, or a menu that is displayed on the screen, that allows the viewer to turn the captioning feature off. However, similar mechanisms are not currently provided for use with data that is delivered over the Internet. Thus, only open captions could be delivered over the Internet. That is, the end users would not be able to control the display of the captioning text, and it would continuously be displayed. Media players, however,

can accommodate closed captioning by allowing end users to close the region 18 in which the captioning data is displayed.

It is, therefore, apparent that there has been provided, in accordance with the present invention, method and system for displaying related components of a media stream that has been transmitted over a computer network . While this invention has been described in conjunction with preferred embodiments thereof, it is evident that many alternatives, modifications, and variations will be apparent to those skilled in the art. Accordingly, it is intended to embrace all such alternatives, modifications and variations that fall within the spirit and broad scope of the appended claims.